Dynamics of Goldring's w-function *

Yong-Gao Chen, Ying Shi, Jie Wu

Dynamics of Goldring's w-function *

Yong-Gao Chen[a], Ying Shi[b] & Jie Wu[c]

[a]School of Mathematical Sciences and Institute of Mathematics,

Nanjing Normal University, Nanjing 210046, P. R. of China

ygchen@njnu.edu.cn

[b]Department of Mathematics, Nanjing Traditional Chinese Medicine University,

Nanjing 210046, P. R. of China

teddy1101@163.com

[c]Institut Elie Cartan Nancy (IECN), CNRS, Nancy-Université, INRIA,

Boulevard des Aiguillettes, B.P. 239, 54506 Vandœuvre-lès-Nancy, France

wujie@iecn.u-nancy.fr

February 23, 2016

Abstract

Let $\mathscr{A}_3$ be the set of all positive integers pqr, where p, q, r are primes such that at least two of them are not equal. Denote by $P(n)$ the largest prime factor of n. For $n = pqr \in \mathscr{A}_3$, define $w(n) := P(p+q)P(p+r)P(q+r)$. In 2006, Wushi Goldring proved that for any $n \in \mathscr{A}_3$, there exists an i such that

*The work of Y.-G. Chen and Y. Shi is supported by the National Natural Science Foundation of China, Grant No 11071121.

$w^i(n) \in \{20, 98, 63, 75\}$, where $w^0(n) = n$ and $w^i(n) = w(w^{i-1}(n))$ $(i \geqslant 1)$. If $w(m) = n$, then m is called a **parent** of n. Let $\mathscr{B}_3$ be the set of all positive integers pq^2 of $\mathscr{A}_3$. In this paper, we study the function w extensively. For example, one of our results is that there exist infinitely many $n \in \mathscr{B}_3$ which have at least $n^{1.1886}$ parents in $\mathscr{B}_3$. Several open problems are posed.

2010 Mathematics Subject Classification: 11A25, 11A41, 37B99

Keywords: w function; primes; dynamics

1 Introduction

Let $\mathscr{P}$ be the set of all (rational) positive primes. For any positive integer n let $P(n)$ denote the largest prime factor of n with the convention $P(1) = 1$. Let

$$\mathscr{A}_3 = \{pqr \mid p,\ q,\ r \in \mathscr{P}\} \setminus \{p^3 \mid p\ \in \mathscr{P}\}.$$

Recently Goldring [4] introduced his w-function on $\mathscr{A}_3$, which is defined by

$$w(n) := P(p+q)P(p+r)P(q+r) \quad (n = pqr \in \mathscr{A}_3),$$

and investigated its dynamics. Here and in the sequel, the letters p, q, r and s denote prime numbers. According to [4, Lemma 2.1], we have $w(n) \in \mathscr{A}_3$ for all $n \in \mathscr{A}_3$. Thus we can consider iteration of w. For every integer $i \geqslant 0$, write

$$w^0(n) := n, \qquad w^i(n) := w(w^{i-1}(n)) \quad (i = 1, 2, \dots).$$

The w-orbit of n is denoted by

$$\mathcal{W}(n) := [n, w(n), \dots, w^i(n), \dots].$$

For example, one can verify that

$$w(20) = 98, \qquad w(98) = 63, \qquad w(63) = 75, \qquad w(75) = 20.$$

Interestingly, Goldring [4, Theorem 1.1] proved that for every $n \in \mathscr{A}_3$, there exists an integer i such that $w^i(n) \in \{20, 98, 63, 75\}$. Denoting by $\mathrm{ind}(n)$ the smallest such integer i. We call $\mathrm{ind}(n)$ *the periodicity index* of n. Goldring [4] proved that

$\mathrm{ind}(n) \leqslant 4(\pi(P(n)) - 3)$, and posed several conjectures related to $w(n)$. Two of them are

Conjecture A ([4, Conjecture 2.9]). We have $\mathrm{ind}(n) = O(\log \pi(P(n)))$.

Conjecture B ([4, Conjecture 2.10]). *There are subsets in $\mathscr{A}_3$ of arbitrarily large periodicity index.*

Chen and Shi [2] proved Conjecture B and $\mathrm{ind}(n) = O((\log P(n))^2)$.

Let $n \in \mathscr{A}_3$ and $\mathscr{S} \subset \mathscr{A}_3$. By a parent of $n \in \mathscr{A}_3$ in $\mathscr{S}$, we mean a positive integer $m \in \mathscr{S}$ such that $w(m) = n$. We also call m a $\mathscr{S}$-parent of n. Let $\mathscr{B}_3 = \{p^2q \,:\, p \neq q,\, p, q \in \mathscr{P}\}$ and $\mathscr{C}_3 = \mathscr{A}_3 \setminus \mathscr{B}_3$. Goldring [4] proved that there exist infinitely many elements of $\mathscr{B}_3$ that have at least seven $\mathscr{B}_3$-parents and posed the following conjecture:

Conjecture C [4, Conjecture 2.16]). *Every element of $\mathscr{A}_3$ (respectively $\mathscr{B}_3$) has infinitely many $\mathscr{C}_3$-parents (respectively $\mathscr{B}_3$).*

Chen and Shi [3] proved that for any positive integer k there exist infinitely many elements of $\mathscr{B}_3$ that have at least k parents in $\mathscr{B}_3$ and there exist infinitely many elements of $\mathscr{B}_3$ that have no parents in $\mathscr{B}_3$.

Define

$$N_{\mathscr{S}}(n) := |\{m \in \mathscr{S} : w(m) = n\}|,$$

Recently Jia [6, Theorem 3] established more precise results:

- There is an element $n = pq^2$ of $\mathscr{B}_3$ with $x < p \leqslant 2x$ and $\sqrt{x}\log x < q \leqslant 2\sqrt{x}\log x$ such that

  $$N_{\mathscr{B}_3}(n) \gg x^{1/2}(\log x)^{-2} \gg n^{1/4}(\log n)^{-5/2},$$

  provided x is sufficiently large.

- There exists an element $n = qr^2$ of $\mathscr{B}_3$ with $q \leqslant 4x$ and $\sqrt{x}\log x < r \leqslant 2\sqrt{x}\log x$ such that

  $$N_{\mathscr{C}_3}(n) \gg x(\log x)^{-4} \gg n^{1/2}(\log n)^{-5},$$

provided x is sufficiently large.

- There is an element $n = qrs$ of $\mathscr{C}_3$ with $q \leqslant 4x$ and $\sqrt{x}\log x < r, s \leqslant 2\sqrt{x}\log x$ such that

$$N_{\mathscr{C}_3}(n) \gg x(\log x)^{-4} \gg n^{1/2}(\log n)^{-5},$$

provided x is sufficiently large.

Here we establish some stronger results.

Theorem 1. (i) *Let $B \geqslant -1$. As $x \to \infty$, we have*

$$\#\{n \leqslant x : n \in \mathscr{B}_3, N_{\mathscr{B}_3}(n) \gg n^{1/3}(\log n)^B\} \gg x^{2/3}(\log x)^{-B-3}. \tag{1.1}$$

(ii) *There exist infinitely many $n \in \mathscr{B}_3$ such that $N_{\mathscr{B}_3}(n) \geqslant n^{1.1886}$.*

Theorem 2. *For any $\varepsilon > 0$, as $x \to \infty$, we have*

$$\#\{n \leqslant x : n \in \mathscr{B}_3, N_{\mathscr{C}_3}(n) \geqslant 1\} \gg_\varepsilon x^{1/3}(\log x)^{-2-\varepsilon}, \tag{1.2}$$

$$\#\{n \leqslant x : n \in \mathscr{B}_3, N_{\mathscr{C}_3}(n) \gg n^{1/2}(\log n)^{-2-\varepsilon}\} \gg \varepsilon \log\log x, \tag{1.3}$$

where the implied constant in (1.3) is absolute.

Theorem 3. *For any $\varepsilon > 0$, as $x \to \infty$, we have*

$$\#\{n \leqslant x : n \in \mathscr{C}_3, N_{\mathscr{C}_3}(n) \gg n^{1/2}(\log n)^{-2-\varepsilon}\} \gg (\varepsilon \log\log x)^2, \tag{1.4}$$

$$\#\{n \leqslant x : n \in \mathscr{C}_3, N_{\mathscr{C}_3}(n) \gg n^{1/2-\varepsilon}\} \gg (\varepsilon \log x)^2, \tag{1.5}$$

where the implied constants are absolute.

In order to improve Goldring's bound to

$$\mathrm{ind}(n) \ll (\log P(n))^2 \qquad (n \in \mathscr{A}_3),$$

Chen & Shi [2] proved their [2, Lemma 4], which is stated as follows: for each $n \in \mathscr{A}_3$ there is a positive integer $i = i(n)$ such that

$$1 \leqslant i \leqslant \log(P(n) + 6) + 2 \quad \text{and} \quad P(w^i(n)) \leqslant \frac{15}{16}P(n) + 6.$$

In Section 5 and 6, we study the distribution of sequences $\{P(w(n))/P(n)\}_{n\in\mathscr{A}_3}$ and $\{w(n)/n\}_{n\in\mathscr{A}_3}$ and establish the following result.

Theorem 4. *We have*

$$\liminf_{\substack{n\to\infty\\ n\in\mathscr{A}_3}} \frac{\log P(w(n))}{\log P(n)} \leqslant 0.2962, \tag{1.6}$$

$$\liminf_{\substack{n\to\infty\\ n\in\mathscr{A}_3}} \frac{\log w(n)}{\log n} \leqslant 0.5924, \tag{1.7}$$

$$\limsup_{\substack{n\to\infty\\ n\in\mathscr{A}_3}} \frac{\log w(n)}{\log n} \geqslant 1.354. \tag{1.8}$$

From Theorem 4, we can deduce the following corollary.

Corollary 5. *We have*

$$\liminf_{\substack{n\to\infty\\ n\in\mathscr{A}_3}} \frac{P(w(n))}{P(n)} = 0 \qquad \text{and} \qquad \limsup_{\substack{n\to\infty\\ n\in\mathscr{A}_3}} \frac{P(w(n))}{P(n)} = 1, \tag{1.9}$$

$$\liminf_{\substack{n\to\infty\\ n\in\mathscr{A}_3}} \frac{w(n)}{n} = 0 \qquad \text{and} \qquad \limsup_{\substack{n\to\infty\\ n\in\mathscr{A}_3}} \frac{w(n)}{n} = \infty. \tag{1.10}$$

Introduce the notation

$$A_3(x) := |\mathscr{A}_3 \cap [1, x]|. \tag{1.11}$$

Theorem 6. *For all $\alpha \in (\frac{1}{2}, 1]$, we have, as $x \to \infty$,*

$$\#\{n \leqslant x : n \in \mathscr{A}_3, P(w(n)) \leqslant \alpha P(n)\} \sim A_3(x). \tag{1.12}$$

When $\alpha \in (0, \frac{1}{2}]$, we have, as $x \to \infty$,

$$\#\{n \leqslant x : n \in \mathscr{A}_3, P(w(n)) \leqslant \alpha P(n)\} \gg \alpha^2 A_3(x), \tag{1.13}$$

where the implied constant is absolute.

In view of Corollary 5, it is natural to raise the following problems.

Problem 7. *Is the set $\{P(w(n))/P(n) : n \in \mathscr{A}_3\}$ dense in $[0, 1]$?*

Problem 8. *Is the set $\{w(n)/n : n \in \mathscr{A}_3\}$ dense in $[0, \infty)$?*

Theorem 6 shows that the density

$$D_{\mathscr{A}_3}(\alpha) := \lim_{x\to\infty} \frac{1}{A_3(x)} \sum_{\substack{n\leqslant x,\, n\in\mathscr{A}_3\\ P(w(n))\leqslant\alpha P(n)}} 1$$

exists for $\frac{1}{2} < \alpha \leqslant 1$ and is equal to 1. Naturally we want to know the other case.

Problem 9. *Does $D_{\mathscr{A}_3}(\alpha)$ exist for $\alpha \in (0, \frac{1}{2}]$? If it exists, what is its value?*

The lower bound (1.1) of Theorem 1 is probably far from optimal. On the other hand, Chen & Shi [3, Theorem 4] proved that there exist infinitely many elements of $\mathscr{B}_3$ which have no $\mathscr{B}_3$-parent. It is very natural to ask

Problem 10. *Does the set $\{n \in \mathscr{B}_3 : n \text{ has no } \mathscr{B}_3\text{-parent}\}$ have zero density?*

In view of [4, Conjecture 2.16], it is interesting to know if there exist elements $n \in \mathscr{B}_3$ such that $N_{\mathscr{B}_3}(n) = \infty$. As Chen & Shi [3, Remark] indicated, this assertion is equivalent to the following conjecture: For any nonzero integer a there are infinitely many primes p such that $P(a+p)$ takes the same value. This conjecture seems very difficult. Here we propose a slightly easier question.

Problem 11. *Let q be a large prime number. Are there infinitely many primes p such that $P(p+2) = q$?*

If the answer to this problem is affirmative, then $N_{\mathscr{B}_3}(2q^2) = \infty$ (since $w(4p) = 2q^2$). In Section 7, we shall give a heuristic argument for Problem 11.

Acknowledgement. This work was begun during Conference on Number Theory at Chern Institute of Mathematics, Nankai University, China, August 23–28, 2009. We would like to thank Chern Institute of Mathematics of Nankai University for agreeable hospitality and Jia for his excellent talk on the dynamics of Goldring's w-function. We also thank the referee for his/her helpful comments.

2 Proof of Theorem 1

First we need to prove a preliminary lemma, which is of independent interest.

Lemma 2.1. *For two coprime positive integers a and d with $1 \leqslant a \leqslant d-1$, denote by $\pi(x, d, a)$ the number of primes $p \leqslant x$ such that $p \equiv a \pmod d$. Let $B_1 \geqslant 0$ and*

$P = Q(\log Q)^{-B_1}$. *Then we have*

$$\sum_{\substack{Q<q\leqslant 2Q \\ \pi(Q^2,q,q-p)\geqslant Q/(8\log Q)}} \sum_{p\leqslant P} 1 \sim \frac{PQ}{\log P \log Q} \sim \frac{Q^2}{(\log Q)^{B_1+2}}, \tag{2.1}$$

$$\sum_{\substack{Q<q\leqslant 2Q \\ N_{\mathscr{B}_3}(pq^2)\geqslant Q/(8\log Q)}} \sum_{p\leqslant P} 1 \sim \frac{PQ}{\log P \log Q} \sim \frac{Q^2}{(\log Q)^{B_1+2}}, \tag{2.2}$$

as $Q \to \infty$.

Proof. According to [7, Theorem 7.12], for $1 \leqslant Q \leqslant x$ and any $A > 0$ we have

$$\sum_{d\leqslant Q} \sum_{\substack{1\leqslant a\leqslant d \\ (a,d)=1}} |E(x,d,a)|^2 = Qx\log Q + O_A\bigg(Qx + \frac{x^2}{(\log x)^A}\bigg), \tag{2.3}$$

where

$$E(x,d,a) := \sum_{\substack{p\leqslant x \\ p\equiv a (\bmod d)}} \log p - \frac{x}{\varphi(d)} \tag{2.4}$$

and $\varphi(q)$ is the Euler function. Here the implied constant depends only on A. Clearly (2.3) with the choice of $x = Q^2$ implies

$$\begin{aligned} \sum_{Q<q\leqslant 2Q} \sum_{p\leqslant P} |E(Q^2,q,q-p)|^2 &\leqslant \sum_{Q<q\leqslant 2Q} \sum_{\substack{1\leqslant a\leqslant q \\ (a,q)=1}} |E(Q^2,q,a)|^2 \\ &\ll_A \frac{Q^4}{(\log Q)^A} \end{aligned} \tag{2.5}$$

for all $Q \geqslant 3$.

Introduce

$$\begin{aligned} S(Q) &:= \{(q,p) : Q < q \leqslant 2Q,\ p \leqslant P\}, \\ E(Q) &:= \big\{(q,p) \in S(Q) : |E(Q^2,q,q-p)| \geqslant Q^2/(2\varphi(q))\big\}. \end{aligned}$$

From the inequality (2.5), we easily deduce

$$|E(Q)|\bigg(\frac{Q^2}{2(2Q-1)}\bigg)^2 \ll_A \frac{Q^4}{(\log Q)^A}.$$

Thus

$$|E(Q)| \ll_A \frac{Q^2}{(\log Q)^A}. \tag{2.6}$$

For all $(q,p) \in S(Q)\smallsetminus E(Q)$, we have

$$|E(Q^2,q,q-p)| < \frac{Q^2}{2\varphi(q)}.$$

From this we deduce

$$\sum_{\substack{p' \leqslant Q^2 \\ p' \equiv q-p \pmod q}} \log p' > \frac{Q^2}{2\varphi(q)}$$

for all $(q,p) \in S(Q)\smallsetminus E(Q)$. Thus

$$\pi(Q^2,q,q-p) \geqslant \frac{Q}{8\log Q}$$

for all $(q,p) \in S(Q)\smallsetminus E(Q)$. Taking $A = B_1 + 3$ in (2.6), we can write

$$\begin{aligned}
\sum_{Q<q\leqslant 2Q} \sum_{p\leqslant P} 1 &= \sum_{\substack{Q<q\leqslant 2Q \\ |E(Q^2,q,q-p)|<Q^2/(2\varphi(q))}} \sum_{p\leqslant P} 1 + O\left(\frac{Q^2}{(\log Q)^{B_1+3}}\right) \\
&\leqslant \sum_{\substack{Q<q\leqslant 2Q \\ \pi(Q^2,q,q-p)\geqslant Q/(8\log Q)}} \sum_{p\leqslant P} 1 + O\left(\frac{Q^2}{(\log Q)^{B_1+3}}\right) \qquad (2.7) \\
&\leqslant \sum_{Q<q\leqslant 2Q} \sum_{p\leqslant P} 1 + O\left(\frac{Q^2}{(\log Q)^{B_1+3}}\right).
\end{aligned}$$

Now the required asymptotic formula (2.1) follows from (2.7) and the prime number theorem.

Now we prove (2.2). For each pair (q,p) with $Q < q \leqslant 2Q, p \leqslant P$ and $\pi(Q^2,q,q-p) \geqslant Q/(8\log Q)$, there are at least $Q/(8\log Q)$ prime numbers $p_i \leqslant Q^2$ such that $p_i \equiv q-p \,(\mathrm{mod}\, q)$. Writting $p_i = \ell_i q + q - p$, we must have $1 \leqslant \ell_i \leqslant q-1$ for $i = 1, \ldots, \lceil Q/(8\log Q)\rceil$. Thus

$$w(p^2 p_i) = pP(p+p_i)^2 = pP((\ell_i+1)q)^2 = pq^2$$

and

$$N_{\mathscr{B}_3}(pq^2) \geqslant \frac{Q}{8\log Q}.$$

From these and (2.1), we deduce that

$$\begin{aligned}
\frac{Q^2}{(\log Q)^{B_1+2}} &\sim \sum_{\substack{Q<q\leqslant 2Q \\ \pi(Q^2,q,q-p)\geqslant Q/(8\log Q)}} \sum_{p\leqslant P} 1 \\
&\leqslant \sum_{\substack{Q<q\leqslant 2Q \\ N_{\mathscr{B}_3}(pq^2)\geqslant Q/(8\log Q)}} \sum_{p\leqslant P} 1 \\
&\leqslant \sum_{Q<q\leqslant 2Q} \sum_{p\leqslant P} 1 \\
&\sim \frac{Q^2}{(\log Q)^{B_1+2}}.
\end{aligned}$$

This is equivalent to (2.2). □

Now we are ready to prove Theorem 1.

The inequality (1.1) is an immediate consequence of (2.2) by writtng $n = pq^2$ and $B_1 = 3B + 3$.

Next we prove the second assertion. According to [1, Theorem 1], for any given prime q, there are two positive constants $c_0 = c_0(q)$ and $x_0 = x_0(q)$ such that

$$\sum_{p\leqslant x,\, P(p+q)\leqslant x^{0.2961}} 1 > \frac{x}{(\log x)^{c_0}} \qquad (x \geqslant x_0).$$

Hence there exists a prime $r \leqslant x^{0.2961}$ such that

$$\sum_{p\leqslant x,\, P(p+q)=r} 1 \geqslant \frac{x}{\pi(x^{0.2961})(\log x)^{c_0}} > \frac{0.2x^{0.7039}}{(\log x)^{c_0-1}} \qquad (x \geqslant x_0).$$

Let $p_1 < \cdots < p_k \leqslant x$ with $P(p_i + q) = r$ $(1 \leqslant i \leqslant k = [0.2x^{0.7039}/(\log x)^{c_0-1}])$. Then $w(p_i q^2) = qr^2$ for $i = 1, 2, \dots, k$ and $qr^2 \leqslant qx^{0.5922} < x^{0.5922} \log x$. Thus for $n_q = qr^2$ we have

$$N_{\mathscr{B}_3}(n_q) \geqslant k > 0.1x^{0.7039}/(\log x)^{c_0-1} > n_q^{1.1886}.$$

Since there are infinitely many prime numbers q, there are infinitely many $n \in \mathscr{B}_3$ such that $N_{\mathscr{C}_3}(n) \geqslant n^{1.1886}$.

3 Proof of Theorem 2

As before, we first prove a preliminary lemma, which is an analogue of Lemma 2.1.

Lemma 3.1. *For two coprime positive integers a and d with $1 \leqslant a \leqslant d-1$, denote by $\pi(x,d,a)$ the number of primes $p \leqslant x$ such that $p \equiv a \,(\operatorname{mod} d)$. Then, for any $B > 3$, we have*

$$\sum_{\substack{Q<q\leqslant 2Q \\ \pi(Q(\log Q)^B, q, q-p) \gg (\log Q)^{B-1}}} \sum_{p\leqslant Q} 1 \sim \frac{Q^2}{(\log Q)^2}, \tag{3.1}$$

$$\sum_{\substack{Q<q\leqslant 2Q \\ \exists\, r \leqslant Q(\log Q)^B \text{ such that } N_{\mathscr{C}_3}(q^2 r)\geqslant 1}} \sum_{p\leqslant Q} 1 \sim \frac{Q^2}{(\log Q)^2}, \tag{3.2}$$

as $Q \to \infty$.

Proof. We shall prove only (3.2), since (3.1) is very similar to (2.1).

By using (3.1), for each (q,p) counted in the left-hand side of (3.1), there are $c(\log Q)^{B-1}$ prime numbers $p_i \leqslant Q(\log Q)^B$ such that $p_i \equiv q-p \,(\operatorname{mod} q)$. Writting $p_i = \ell_i q + q - p$, we must have $1 \leqslant \ell_i \leqslant (\log Q)^B < Q \leqslant q-1$ for $i = 1, \ldots, [c(\log Q)^{B-1}]$. Thus

$$w(pp_ip_j) = P(p+p_i)P(p+p_j)P(p_i+p_j) = q^2 P(p_i+p_j).$$

Clearly $p \neq p_i$ and $P(p_i+p_j) \neq q$ for all $1 \leqslant i,j \leqslant [c(\log Q)^{B-1}]$. Taking $r = P(p_i+p_j)$, we have

$$N_{\mathscr{C}_3}(q^2 r) \geqslant 1.$$

This and (3.1) imply our required result. $\square$

Now we are ready to prove Theorem 2.

From (3.2), we easily deduce that

$$\sum_{\substack{Q<q\leqslant 2Q \\ \exists\, r \leqslant Q(\log Q)^B \text{ such that } N_{\mathscr{C}_3}(q^2 r)\geqslant 1}} 1 \gg \frac{Q}{\log Q}$$

and

$$\sum_{\substack{n\leqslant x,\, n\in\mathscr{B}_3 \\ N_{\mathscr{C}_3}(n)\geqslant 1}} 1 \geqslant \sum_{\substack{0.5x^{1/3}(\log x)^{-B/3} < q \leqslant x^{1/3}(\log x)^{-B/3} \\ \exists\, r\leqslant x^{1/3}(\log x)^{2B/3} \text{ such that } N_{\mathscr{C}_3}(q^2 r)\geqslant 1}} 1 \gg \frac{x^{1/3}}{(\log x)^{(B+3)/3}}.$$

This proves the inequality (1.2) by taking $B = 3\varepsilon + 3$.

Next we shall prove the estimate (1.3). The method of proof is the same as in [6, Theorem 2]. The new ingredient is that we add a process of summation, which allows us to get the lower bound $\gg \log\log x$.

For $k \geqslant 1$ and $2^{k+1} \leqslant x^{1/2}/(\log x)^7$, we consider the sum

$$\begin{aligned}\sigma_k &:= \sum_{\substack{x<p_1\leqslant 2x \quad x<p_2\leqslant 2x \quad x<p_3\leqslant 2x \\ 2^k x^{1/2}<P(p_1+p_2)=P(p_1+p_3)\leqslant 2^{k+1}x^{1/2}}} \log p_2 \log p_3 \\ &= \sum_{x<p_1\leqslant 2x} \sum_{2^k x^{1/2}<r\leqslant 2^{k+1}x^{1/2}} \Bigg(\sum_{\substack{x<p_2\leqslant 2x \\ P(p_1+p_2)=r}} \log p_2 \Bigg)^2.\end{aligned} \tag{3.3}$$

Since for all $x < p_1 \leqslant 2x$ and $2^k x^{1/2} < r \leqslant 2^{k+1}x^{1/2}$, we have

$$x < p_2 \leqslant 2x \text{ and } P(p_1+p_2) = r \quad \Leftrightarrow \quad x < p_2 \leqslant 2x \text{ and } p_2 \equiv -p_1 (\operatorname{mod} r),$$

we can write

$$\sum_{\substack{x<p_2\leqslant 2x \\ P(p_1+p_2)=r}} \log p_2 = \frac{x}{\varphi(r)} + E^*(x, r, -p_1),$$

where

$$E^*(x, r, -p_1) = E(2x, r, -p_1) - E(x, r, -p_1) \tag{3.4}$$

and $E(x, d, a)$ is defined as in (2.4). Thus

$$\sigma_k = \sigma_k^{(1)} + 2\sigma_k^{(2)} + \sigma_k^{(3)}, \tag{3.5}$$

where

$$\begin{aligned}\sigma_k^{(1)} &:= \sum_{x<p_1\leqslant 2x} \sum_{2^k x^{1/2}<r\leqslant 2^{k+1}x^{1/2}} \left(\frac{x}{\varphi(r)}\right)^2, \\ \sigma_k^{(2)} &:= \sum_{x<p_1\leqslant 2x} \sum_{2^k x^{1/2}<r\leqslant 2^{k+1}x^{1/2}} \frac{x}{\varphi(r)} E^*(x, r, -p_1), \\ \sigma_k^{(3)} &:= \sum_{x<p_1\leqslant 2x} \sum_{2^k x^{1/2}<r\leqslant 2^{k+1}x^{1/2}} |E^*(x, r, -p_1)|^2.\end{aligned}$$

By the prime number theorem, it is easy to see that

$$\sigma_k^{(1)} \geqslant \left(\frac{x}{2^{k+1}x^{1/2}}\right)^2 \sum_{x<p_1\leqslant 2x} \sum_{2^k x^{1/2}<r\leqslant 2^{k+1}x^{1/2}} 1 \gg \frac{2^{-k}x^{5/2}}{(\log x)^2}\left\{1 + O\left(\frac{1}{\log x}\right)\right\} \tag{3.6}$$

uniformly for

$$k \geqslant 1 \qquad \text{and} \qquad 2^k \leqslant x^{1/2}/(\log x)^7. \tag{3.7}$$

By the simple inequality

$$|E^*(x,r,-p_1)|^2 \leqslant 2\big(|E(2x,r,-p_1)|^2 + |E(x,r,-p_1)|^2\big)$$

and formula (2.3) with $A=4$, we have

$$\begin{aligned}\sigma_k^{(3)} &= \sum_{2^k x^{1/2}<r\leqslant 2^{k+1}x^{1/2}} \sum_{1\leqslant a<r} \sum_{\substack{x<p_1\leqslant 2x\\ p_1\equiv -a \,(\mathrm{mod}\, r)}} |E^*(x,r,a)|^2 \\ &\ll \sum_{2^k x^{1/2}<r\leqslant 2^{k+1}x^{1/2}} \sum_{1\leqslant a<r} \frac{x}{r}|E^*(x,r,a)|^2 \\ &\ll \frac{2^{-k}x^{5/2}}{(\log x)^4},\end{aligned} \tag{3.8}$$

uniformly for x and k satisfying (3.7).

By the Cauchy-Schwarz inequality and the following simple bound

$$\sum_{x<p_1\leqslant 2x} \sum_{2^k x^{1/2}<r\leqslant 2^{k+1}x^{1/2}} 1 \ll \frac{2^k x^{3/2}}{(\log x)^2}$$

uniformly for x and k satisfying (3.7), it follows that

$$\begin{aligned}\big|\sigma_k^{(2)}\big|^2 &\ll \bigg(2^{-k}x^{1/2} \sum_{x<p_1\leqslant 2x} \sum_{2^k x^{1/2}<r\leqslant 2^{k+1}x^{1/2}} |E^*(x,r,-p_1)|\bigg)^2 \\ &\ll \frac{2^{-k}x^{5/2}}{(\log x)^2} \sum_{x<p_1\leqslant 2x} \sum_{2^k x^{1/2}<r\leqslant 2^{k+1}x^{1/2}} |E^*(x,r,-p_1)|^2 \\ &= \frac{2^{-k}x^{5/2}}{(\log x)^2}\sigma_k^{(3)} \ll \frac{2^{-2k}x^5}{(\log x)^6}\end{aligned} \tag{3.9}$$

uniformly for x and k satisfying (3.7).

Inserting (3.6)-(3.9) into (3.5), we find that

$$\sigma_k \gg \frac{2^{-k}x^{5/2}}{(\log x)^2}\bigg\{1 + O\bigg(\frac{1}{\log x}\bigg)\bigg\} \tag{3.10}$$

uniformly for x and k satisfying (3.7).

Since

$$\sum_{x<p_1\leqslant 2x} \sum_{\substack{x<p_2\leqslant 2x\\ p_1p_2p_3\notin\mathscr{C}_3}} \sum_{x<p_3\leqslant 2x} \log p_2 \log p_3 \ll x^2 \log x,$$

the inequality (3.10) yields immediately

$$\sigma_k^* := \sum_{\substack{x<p_1\leqslant 2x \\ 2^k x^{1/2}<P(p_1+p_2)=P(p_1+p_3)\leqslant 2^{k+1}x^{1/2} \\ p_1p_2p_3\in\mathscr{C}_3}} \sum_{x<p_2\leqslant 2x} \sum_{x<p_3\leqslant 2x} \log p_2 \log p_3 \gg \frac{2^{-k}x^{5/2}}{(\log x)^2}\bigg\{1+O\bigg(\frac{1}{\log x}\bigg)\bigg\} \tag{3.11}$$

uniformly for x and k satisfying (3.7). On the other hand, we can write

$$\sigma_k^* = \sum_{q\leqslant 4x} \sum_{\substack{2^k x^{1/2}<r\leqslant 2^{k+1}x^{1/2} \\ qr^2\in\mathscr{B}_3}} \sigma_k(q,r)$$

with

$$\sigma_k(q,r) := \sum_{\substack{x<p_1\leqslant 2x \\ P(p_2+p_3)=q,\, P(p_1+p_2)=P(p_1+p_3)=r \\ p_1p_2p_3\in\mathscr{C}_3}} \sum_{x<p_2\leqslant 2x} \sum_{x<p_3\leqslant 2x} \log p_2 \log p_3.$$

Since we have

$$\sum_{q\leqslant 4x} \sum_{\substack{2^k x^{1/2}<r\leqslant 2^{k+1}x^{1/2} \\ qr^2\in\mathscr{B}_3}} 1 \ll \frac{2^k x^{3/2}}{(\log x)^2}$$

uniformly for x and k satisfying (3.7), the inequality (3.11) guarantees that there is at least a couple (q,r) satisfying $q\leqslant 4x$ and $2^k x^{1/2}<r\leqslant 2^{k+1}x^{1/2}$ such that

$$\sigma_k(q,r) \gg \frac{x}{2^{2k}} \gg \frac{x}{(\log x)^{2\varepsilon}} \qquad (2^k\leqslant (\log x)^{\varepsilon}).$$

In other words, for each integer k with $2^k\leqslant(\log x)^{\varepsilon}$, there is at least a couple (q,r) with $q\leqslant 4x$ and $2^k x^{1/2}<r\leqslant 2^{k+1}x^{1/2}$ such that $N_{\mathscr{C}_3}(qr^2)\gg x/(\log x)^{2+2\varepsilon}$. Thus

$$\begin{aligned}\sum_{\substack{n\leqslant 4x^2(\log x)^{2\varepsilon},\, n\in\mathscr{B}_3 \\ N_{\mathscr{C}_3}(n)\gg x/(\log x)^{2+2\varepsilon}}} 1 &\geqslant \sum_{\substack{qr^2\leqslant 4x^2(\log x)^{2\varepsilon},\, qr^2\in\mathscr{B}_3 \\ q\leqslant 4x,\, 2x^{1/2}\leqslant r\leqslant x^{1/2}(\log x)^{\varepsilon} \\ N_{\mathscr{C}_3}(qr^2)\gg x/(\log x)^{2+2\varepsilon}}} 1 \\ &\geqslant \sum_{k\leqslant\varepsilon\log\log x} \sum_{\substack{qr^2\leqslant 4x^2(\log x)^{2\varepsilon},\, qr^2\in\mathscr{B}_3 \\ q\leqslant 4x,\, 2^k x^{1/2}\leqslant r\leqslant 2^{k+1}x^{1/2} \\ N_{\mathscr{C}_3}(qr^2)\gg x/(\log x)^{2+2\varepsilon}}} 1 \\ &\geqslant \varepsilon\log\log x - 1.\end{aligned}$$

This implies (1.3) by replacing $4x^2(\log x)^{2\varepsilon}$ with x.

4 Proof of Theorem 3

Since the proof is similar to (1.3), we shall point out the principal lines only.

For $2^{j+1} \leqslant x^{1/2}/(\log x)^7$ and $2^{k+1} \leqslant x^{1/2}/(\log x)^7$ with $j \geqslant 1$ and $k \geqslant 1$, consider the sum

$$\sigma_{j,k} := \sum_{\substack{x<p_1\leqslant 2x \\ 2^j x^{1/2}<P(p_1+p_2)\leqslant 2^{j+1}x^{1/2},\ 2^k x^{1/2}<P(p_1+p_3)\leqslant 2^{k+1}x^{1/2}}} \sum_{x<p_2\leqslant 2x} \sum_{x<p_3\leqslant 2x} \log p_2 \log p_3. \tag{4.1}$$

Similar to (3.5), we have

$$\sigma_{j,k} = \sigma_{j,k}^{(1)} + \sigma_{j,k}^{(2)} + \sigma_{j,k}^{(3)} + \sigma_{j,k}^{(4)}, \tag{4.2}$$

where

$$\begin{aligned}
\sigma_{j,k}^{(1)} &:= \sum_{x<p_1\leqslant 2x} \sum_{2^j x^{1/2}<q\leqslant 2^{j+1}x^{1/2}} \frac{x}{\varphi(q)} \sum_{2^k x^{1/2}<r\leqslant 2^{k+1}x^{1/2}} \frac{x}{\varphi(r)}, \\
\sigma_{j,k}^{(2)} &:= \sum_{x<p_1\leqslant 2x} \sum_{2^j x^{1/2}<q\leqslant 2^{j+1}x^{1/2}} \frac{x}{\varphi(q)} \sum_{2^k x^{1/2}<r\leqslant 2^{k+1}x^{1/2}} E^*(x,r,-p_1), \\
\sigma_{j,k}^{(3)} &:= \sum_{x<p_1\leqslant 2x} \sum_{2^j x^{1/2}<q\leqslant 2^{j+1}x^{1/2}} E^*(x,q,-p_1) \sum_{2^k x^{1/2}<r\leqslant 2^{k+1}x^{1/2}} \frac{x}{\varphi(r)}, \\
\sigma_{j,k}^{(4)} &:= \sum_{x<p_1\leqslant 2x} \sum_{2^j x^{1/2}<q\leqslant 2^{j+1}x^{1/2}} E^*(x,q,-p_1) \sum_{2^k x^{1/2}<r\leqslant 2^{k+1}x^{1/2}} E^*(x,r,-p_1),
\end{aligned}$$

and $E^*(x,r,-p_1)$ is defined in (3.4).

By the prime number theorem, it is easy to see that

$$\sigma_{j,k}^{(1)} \geqslant (\log 2)^2 \frac{x^3}{(\log x)^3} \left\{1 + O\left(\frac{1}{\log x}\right)\right\} \tag{4.3}$$

uniformly for

$$j,k \geqslant 1, \qquad 2^j \leqslant x^{1/2}/(\log x)^7 \quad \text{and} \quad 2^k \leqslant x^{1/2}/(\log x)^7. \tag{4.4}$$

Similar to $\sigma_k^{(2)}$, with the help of (2.3) with $A = 6$ we easily prove that

$$\sigma_{j,k}^{(i)} \ll \frac{x^3}{(\log x)^4} \qquad (i = 2, 3) \tag{4.5}$$

uniformly for x, j and k satisfying (4.4).

By applying the Cauchy-Schwarz inequality two times, it follows that

$$\begin{aligned}\big|\sigma_{j,k}^{(4)}\big|^2 &\leqslant \sum_{x<p_1\leqslant 2x}\ \sum_{2^jx^{1/2}<q\leqslant 2^{j+1}x^{1/2}} |E^*(x,q,-p_1)|^2\times \\ &\quad\times \sum_{x<p_1'\leqslant 2x}\ \sum_{2^jx^{1/2}<q'\leqslant 2^{j+1}x^{1/2}} \bigg(\sum_{2^kx^{1/2}<r\leqslant 2^{k+1}x^{1/2}} E^*(x,r,-p_1')\bigg)^2 \\ &\leqslant \sum_{2^jx^{1/2}<q'\leqslant 2^{j+1}x^{1/2}}\ \sum_{2^kx^{1/2}<r'\leqslant 2^{k+1}x^{1/2}} \sigma_j^{(3)}\sigma_k^{(3)}.\end{aligned}$$

Thus (3.8) implies that

$$\big|\sigma_{j,k}^{(4)}\big|^2 \ll \frac{2^jx^{1/2}}{\log x}\frac{2^kx^{1/2}}{\log x}\frac{2^{-j}x^{5/2}}{(\log x)^4}\frac{2^{-k}x^{5/2}}{(\log x)^4} \ll \frac{x^6}{(\log x)^{10}} \tag{4.6}$$

uniformly for x, j and k satisfying (4.4).

Inserting (4.3)-(4.6) into (4.2), we find that

$$\sigma_{j,k} \geqslant (\log 2)^2\frac{x^3}{(\log x)^3}\bigg\{1+O\bigg(\frac{1}{\log x}\bigg)\bigg\} \tag{4.7}$$

uniformly for x, j and k satisfying (4.4).

In order to remove p_1, p_2, p_3 in $\sigma_{j,k}$ such that $w(p_1p_2p_3) \notin \mathscr{C}_3$, we consider the subsums of $\sigma_{j,k}$:

$$\sigma_{j,k}' := \sum_{\substack{x<p_1\leqslant 2x\ x<p_2\leqslant 2x\ x<p_3\leqslant 2x \\ 2^jx^{1/2}<P(p_1+p_2)\leqslant 2^{j+1}x^{1/2},\ 2^kx^{1/2}<P(p_1+p_3)\leqslant 2^{k+1}x^{1/2} \\ P(p_1+p_2)=P(p_1+p_3)}} \log p_2\log p_3,$$

$$\sigma_{j,k}'' := \sum_{\substack{x<p_1\leqslant 2x\ x<p_2\leqslant 2x\ x<p_3\leqslant 2x \\ 2^jx^{1/2}<P(p_1+p_2)\leqslant 2^{j+1}x^{1/2},\ 2^kx^{1/2}<P(p_1+p_3)\leqslant 2^{k+1}x^{1/2} \\ P(p_1+p_2)=P(p_2+p_3)}} \log p_2\log p_3,$$

$$\sigma_{j,k}''' := \sum_{\substack{x<p_1\leqslant 2x\ x<p_2\leqslant 2x\ x<p_3\leqslant 2x \\ 2^jx^{1/2}<P(p_1+p_2)\leqslant 2^{j+1}x^{1/2},\ 2^kx^{1/2}<P(p_1+p_3)\leqslant 2^{k+1}x^{1/2} \\ P(p_1+p_3)=P(p_2+p_3)}} \log p_2\log p_3.$$

For $\sigma'_{j,k}$, we must have $j = k$ and

$$\begin{aligned}
\sigma'_{j,k} &= \sum_{2^j x^{1/2} < q \leqslant 2^{j+1} x^{1/2}} \sum_{x < p_1 \leqslant 2x} \sum_{\substack{x < p_2 \leqslant 2x \\ P(p_1+p_2)=q}} \sum_{\substack{x < p_3 \leqslant 2x \\ P(p_1+p_3)=q}} \log p_2 \log p_3 \\
&\leqslant 4(\log x)^2 \sum_{2^j x^{1/2} < q \leqslant 2^{j+1} x^{1/2}} \sum_{x < p_1 \leqslant 2x} \sum_{\substack{x < p_2 \leqslant 2x \\ p_2 \equiv -p_1 \ (\mathrm{mod}\, q)}} \sum_{\substack{x < p_3 \leqslant 2x \\ p_3 \equiv -p_1 \ (\mathrm{mod}\, q)}} 1 \\
&\leqslant 16(\log x)^2 \sum_{2^j x^{1/2} < q \leqslant 2^{j+1} x^{1/2}} \sum_{x < p_1 \leqslant 2x} \frac{x^2}{q^2} \\
&\leqslant 20 x^{5/2} (\log x)^2.
\end{aligned}$$

The same bound holds for $\sigma''_{j,k}$ and $\sigma'''_{j,k}$. These estimates allow us to remove the p_1, p_2, p_3 in $\sigma_{j,k}$ such that $w(p_1p_2p_3) \notin \mathscr{C}_3$:

$$\begin{aligned}
\sigma^*_{j,k} &:= \sum_{\substack{x < p_1 \leqslant 2x \ x < p_2 \leqslant 2x \ x < p_3 \leqslant 2x \\ 2^j x^{1/2} < P(p_1+p_2) \leqslant 2^{j+1} x^{1/2},\ 2^k x^{1/2} < P(p_1+p_3) \leqslant 2^{k+1} x^{1/2} \\ w(p_1p_2p_3) \in \mathscr{C}_3}} \log p_2 \log p_3 \\
&\geqslant (\log 2)^2 \frac{x^3}{(\log x)^3} \left\{ 1 + O\left(\frac{1}{\log x} \right) \right\}
\end{aligned} \tag{4.8}$$

uniformly for x, j and k satisfying (4.4). On the other hand, we can write

$$\sigma^*_{j,k} = \sum_{q \leqslant 4x} \sum_{\substack{2^j x^{1/2} < r \leqslant 2^{j+1} x^{1/2} \quad 2^k x^{1/2} < s \leqslant 2^{k+1} x^{1/2} \\ qrs \in \mathscr{C}_3}} \sigma_{j,k}(q, r, s)$$

with

$$\sigma_{j,k}(q, r, s) := \sum_{\substack{x < p_1 \leqslant 2x \ x < p_2 \leqslant 2x \ x < p_3 \leqslant 2x \\ P(p_2+p_3)=q,\ P(p_1+p_2)=r,\ P(p_1+p_3)=s}} \log p_2 \log p_3.$$

Since we have

$$\sum_{q \leqslant 4x} \sum_{\substack{2^j x^{1/2} < r \leqslant 2^{j+1} x^{1/2} \quad 2^k x^{1/2} < s \leqslant 2^{k+1} x^{1/2} \\ qrs \in \mathscr{C}_3}} 1 \ll \frac{2^{j+k} x^2}{(\log x)^3}$$

uniformly for x, j and k satisfying (4.4). The inequality (4.8) guarantees that there is at least a triple (q, r, s) with $q \leqslant 4x$, $2^j x^{1/2} < r \leqslant 2^{j+1} x^{1/2}$ and $2^k x^{1/2} < s \leqslant 2^{k+1} x^{1/2}$ such that

$$\sigma_{j,k}(q, r, s) \gg \frac{x}{2^{j+k}} \gg \begin{cases} x/(\log x)^{2\varepsilon} & \text{for } 2^j \leqslant (\log x)^{\varepsilon} \text{ and } 2^k \leqslant (\log x)^{\varepsilon}, \\ x^{1-2\varepsilon} & \text{for } 2^j \leqslant x^{\varepsilon} \text{ and } 2^k \leqslant x^{\varepsilon}. \end{cases}$$

In other words, for each couple (j,k) with $2^j \leqslant (\log x)^\varepsilon$ and $2^k \leqslant (\log x)^\varepsilon$ (resp. $2^j \leqslant x^\varepsilon$ and $2^k \leqslant x^\varepsilon$), there is at least a triple (q,r,s) with $q \leqslant 4x$, $2^j x^{1/2} < r \leqslant 2^{j+1}x^{1/2}$ and $2^k x^{1/2} < s \leqslant 2^{k+1}x^{1/2}$ such that $N_{\mathscr{C}_3}(qrs) \gg x/(\log x)^{2+2\varepsilon}$ (resp. $x^{1-3\varepsilon}$). Thus

$$\begin{aligned}\sum_{\substack{n\leqslant 4x^2(\log x)^{2\varepsilon},\, n\in\mathscr{C}_3\\ N_{\mathscr{C}_3}(n)\gg x/(\log x)^{2+2\varepsilon}}} 1 &\geqslant \sum_{\substack{qrs\leqslant 4x^2(\log x)^{2\varepsilon},\, qrs\in\mathscr{C}_3\\ q\leqslant 4x,\, 2x^{1/2}\leqslant r\leqslant x^{1/2}(\log x)^{\varepsilon},\, 2x^{1/2}\leqslant s\leqslant x^{1/2}(\log x)^{\varepsilon}\\ N_{\mathscr{C}_3}(qrs)\gg x/(\log x)^{2+2\varepsilon}}} 1\\ &\geqslant \sum_{j\leqslant \varepsilon\log\log x}\ \sum_{k\leqslant\varepsilon\log\log x}\ \sum_{\substack{qrs\leqslant 4x^2(\log x)^{2\varepsilon},\, qrs\in\mathscr{C}_3\\ q\leqslant 4x,\, 2^jx^{1/2}\leqslant r\leqslant 2^{j+1}x^{1/2},\, 2^kx^{1/2}\leqslant s\leqslant 2^{k+1}x^{1/2}\\ N_{\mathscr{C}_3}(qrs)\gg x/(\log x)^{2+2\varepsilon}}} 1\\ &\geqslant (\varepsilon\log\log x)^2 - 1.\end{aligned}$$

This implies (1.4) by replacing $4x^2(\log x)^{2\varepsilon}$ with x. The inequality (1.5) can proved in the same way.

5 Proofs of Theorem 4 and Corollary 5

According to [1, Theorem 1], there are two positive constants c_1 and x_0 such that

$$\sum_{p\leqslant x,\, P(p+2)\leqslant x^{0.2961}} 1 \gg \frac{x}{(\log x)^{c_1}} \qquad (x\geqslant x_0). \tag{5.1}$$

Since

$$\begin{aligned}\sum_{p\leqslant x,\, P(p+2)\leqslant x^{0.2961}} 1 &= \sum_{x^{0.2961/0.2962}<p\leqslant x,\, P(p+2)\leqslant x^{0.2961}} 1 + O\big(x^{0.2961/0.2962}\big)\\ &\leqslant \sum_{p\leqslant x,\, P(p+2)\leqslant p^{0.2962}} 1 + O\big(x^{0.2961/0.2962}\big),\end{aligned}$$

the inequality (5.1) implies

$$\sum_{p\leqslant x,\, P(p+2)\leqslant p^{0.2962}} 1 \gg \frac{x}{(\log x)^{c_1}} \qquad (x\geqslant x_0).$$

For such p, we have

$$\begin{aligned}\frac{\log P(w(4p))}{\log P(4p)} &= \frac{\log P(p+2)}{\log p} \leqslant 0.2962,\\ \frac{\log w(4p)}{\log(4p)} &= \frac{\log(2P(p+2)^2)}{\log(4p)} \leqslant \frac{0.5924\log p + \log 2}{\log(4p)}.\end{aligned}$$

These imply (1.6) and (1.7).

By applying [1, Theorem 2], there are infinitely many prime numbers p_i with $P(p_i+2) \geqslant p_i^{0.677}$. Thus

$$\frac{\log w(4p_i)}{\log(4p_i)} = \frac{\log(2P(p_i+2)^2)}{\log(4p_i)} \geqslant \frac{1.354\log p_i + \log 2}{\log p_i + \log 4}.$$

Hence (1.8) follows immediately. This proves Theorem 4.

Clearly Theorem 4 implies Corollary 5 except for the second formula of (1.9). In order to prove it, we first notice that $P(w(n)) \leqslant P(n)+2$ for all $n \in \mathscr{A}_3$. Thus

$$\limsup_{\substack{n\to\infty \\ n\in\mathscr{A}_3}} \frac{P(w(n))}{P(n)} \leqslant 1. \tag{5.2}$$

On the other hand, according to the Green-Tao Theorem on arithmetic progressions in primes [5], for any integer $k \geqslant 2$, there are two positive integers $a = a(k)$ and $d = d(k)$ such that $a + id$ is prime for $i = 0, \dots, 2^k$. Put

$$p := a + 2^k d, \qquad q := a + (2^k - 2)d, \qquad r := a$$

and

$$\begin{aligned} p' &:= \frac{p+q}{2} = a + (2^k-1)d, \\ q' &:= \frac{p+r}{2} = a + 2^{k-1}d, \\ r' &:= \frac{q+r}{2} = a + (2^{k-1}-1)d. \end{aligned}$$

Obviously $w(pqr) = p'q'r'$ and $P(w(pqr)) = p'$. Thus

$$\lim_{k\to\infty} \frac{P(w(pqr))}{P(pqr)} = \lim_{k\to\infty} \frac{p'}{p} = 1. \tag{5.3}$$

Now the second relation of (1.9) follows from (5.2) and (5.3).

6 Proof of Theorem 6

We need to establish two preliminary lemmas.

Lemma 6.1. *With the notation of* (1.11) ($B_3(x)$ *and* $C_3(x)$ *can be defined similarly*), *we have, for* $x \to \infty$,

$$A_3(x) \sim \frac{x}{2\log x}(\log\log x)^2, \tag{6.1}$$

$$B_3(x) \sim B_3 \frac{x}{\log x} \quad \text{with} \quad B_3 := \sum_p \frac{1}{p^2}, \tag{6.2}$$

$$C_3(x) \sim \frac{x}{2\log x}(\log\log x)^2. \tag{6.3}$$

Proof. By the prime number theorem, we can write

$$B_3(x) = \sum_{p^2 q \leqslant x,\, p \neq q} 1 \sim \sum_{p \leqslant (x/2)^{1/2}} \frac{x}{p^2 \log(x/p^2)}.$$

On the other hand, we have

$$\begin{aligned}
\sum_{\log x < p \leqslant (x/2)^{1/2}} \frac{x}{p^2 \log(x/p^2)} &\ll \sum_{p > \log x} \frac{x}{p^2} \\
&\ll \frac{x}{(\log x)\log\log x}, \\
\sum_{p \leqslant \log x} \frac{x}{p^2 \log(x/p^2)} &\sim \frac{x}{\log x} \sum_{p \leqslant \log x} \frac{1}{p^2} \\
&\sim B_3 \frac{x}{\log x}.
\end{aligned}$$

Inserting these into the preceding formula, we get (6.2).

According to a classic result of Landau (see [8, Chapter II.6]), we have

$$\sum_{\substack{n \leqslant x \\ \Omega(n)=3}} 1 \sim \frac{x}{2\log x}(\log\log x)^2 \qquad (x \to \infty).$$

From this and (6.2), we immediately deduce (6.3), since

$$C_3(x) = \sum_{\substack{n \leqslant x \\ \Omega(n)=3}} 1 - \sum_{\substack{n \leqslant x \\ \Omega(n)=3,\, \omega(n)=2}} 1 - \sum_{\substack{n \leqslant x \\ \Omega(n)=3,\, \omega(n)=1}} 1.$$

Finally (6.1) follows from (6.2) and (6.3), since $A_3(x) = B_3(x) + C_3(x)$. $\square$

Lemma 6.2. *For $k \in \mathbb{N}$, define*

$$\mathscr{C}_3(k) := \{pqr : r < q < p,\ q \equiv r \,(\mathrm{mod}\, 2^k),\ p \equiv -r \,(\mathrm{mod}\, 2^k)\}.$$

For any $\varepsilon \in (0, \frac{1}{2^{k+1}})$, we have, as $x \to \infty$,

$$\sum_{\substack{n \leqslant x,\, n \in \mathscr{C}_3(k) \\ P(w(n)) \leqslant (1+\varepsilon)2^{-k} P(n)}} 1 \sim \sum_{n \leqslant x,\, n \in \mathscr{C}_3(k)} 1. \tag{6.4}$$

Further we have

$$\sum_{n \leqslant x,\, n \in \mathscr{C}_3(1)} 1 \sim \frac{x}{2 \log x} (\log \log x)^2, \tag{6.5}$$

$$\sum_{n \leqslant x,\, n \in \mathscr{C}_3(k)} 1 \gg \frac{A_3(x)}{2^{2k}}, \tag{6.6}$$

where the implied constant in the $\gg$-symbol is absolute.

Proof. For any $\varepsilon \in (0, \frac{1}{2^{k+1}})$, we have

$$\begin{aligned}
\sum_{r \leqslant x^{1/3}} \sum_{q \leqslant (x/r)^{1/2}} \sum_{p \leqslant q/\varepsilon} 1 &\ll_\varepsilon \sum_{r \leqslant x^{1/3}} \sum_{q \leqslant (x/r)^{1/2}} \frac{q}{\log q} \\
&\ll_\varepsilon \sum_{r \leqslant x^{1/3}} \frac{x}{r \log^2 (x/r)} \\
&\ll_\varepsilon \frac{x \log \log x}{(\log x)^2}.
\end{aligned}$$

Thus we can write

$$\begin{aligned}
\sum_{n \leqslant x,\, n \in \mathscr{C}_3(k)} 1 &= \sum_{r \leqslant x^{1/3}} \sum_{\substack{r < q \leqslant (x/r)^{1/2} \\ q \equiv r (\mathrm{mod}\, 2^k)}} \sum_{\substack{q < p \leqslant x/(qr) \\ p \equiv -r (\mathrm{mod}\, 2^k)}} 1 \\
&= \sum_{r \leqslant x^{1/3}} \sum_{\substack{r < q \leqslant (x/r)^{1/2} \\ q \equiv r (\mathrm{mod}\, 2^k)}} \sum_{\substack{q/\varepsilon < p \leqslant x/(qr) \\ p \equiv -r (\mathrm{mod}\, 2^k)}} 1 + O_\varepsilon\bigg(\frac{x \log \log x}{(\log x)^2} \bigg).
\end{aligned}$$

For each (p, q, r) counted in the last triple sums, we have obviously that

$$\begin{aligned}
P(p+q) &\leqslant \frac{p+q}{2^k} \leqslant \frac{1+\varepsilon}{2^k} p, \\
P(p+r) &\leqslant \frac{p+r}{2^k} \leqslant \frac{1+\varepsilon}{2^k} p, \\
P(q+r) &\leqslant 2q \leqslant 2\varepsilon p \leqslant \frac{1+\varepsilon}{2^k} p.
\end{aligned}$$

Thus

$$P(w(pqr)) = \max\{P(p+q), P(p+r), P(q+r)\} \leqslant (1+\varepsilon)2^{-k}p$$

and

$$\begin{aligned}
\sum_{n\leqslant x,\, n\in\mathscr{C}_3(k)} 1 &\leqslant \sum_{r\leqslant x^{1/3}} \sum_{r<q\leqslant (x/r)^{1/2}} \sum_{\substack{q/\varepsilon<p\leqslant x/(qr) \\ P(w(pqr))\leqslant(1+\varepsilon)2^{-k}P(pqr)}} 1 + O_\varepsilon\biggl(\frac{x\log\log x}{(\log x)^2}\biggr) \\
&\leqslant \sum_{r\leqslant x^{1/3}} \sum_{r<q\leqslant (x/r)^{1/2}} \sum_{\substack{q<p\leqslant x/(qr) \\ P(w(pqr))\leqslant(1+\varepsilon)2^{-k}P(pqr)}} 1 + O_\varepsilon\biggl(\frac{x\log\log x}{(\log x)^2}\biggr) \\
&= \sum_{\substack{n\leqslant x,\, n\in\mathscr{C}_3(k) \\ P(w(n))\leqslant(1+\varepsilon)2^{-k}P(n)}} 1 + O_\varepsilon\biggl(\frac{x\log\log x}{(\log x)^2}\biggr).
\end{aligned}$$

This implies (6.4).

The asymptotic formula (6.5) is an immediate consequence of (6.3), since

$$\begin{aligned}
\sum_{n\leqslant x,\, n\in\mathscr{C}_3} 1 &= \sum_{r\leqslant x^{1/3}} \sum_{r<q\leqslant (x/r)^{1/2}} \sum_{q<p\leqslant x/(qr)} 1 \\
&= \sum_{3\leqslant r\leqslant x^{1/3}} \sum_{r<q\leqslant (x/r)^{1/2}} \sum_{q<p\leqslant x/(qr)} 1 + O\biggl(\frac{x\log\log x}{\log x}\biggr) \\
&= \sum_{n\leqslant x,\, n\in\mathscr{C}_3(1)} 1 + O\biggl(\frac{x\log\log x}{\log x}\biggr).
\end{aligned}$$

In order to show the lower bound of (6.6), we begin by the following trivial inequality

$$\sum_{\substack{n\leqslant x \\ n\in\mathscr{C}_3(k)}} 1 \geqslant \sum_{r\leqslant x^{1/6}} \sum_{\substack{r<q\leqslant x^{1/3} \\ q\equiv r (\operatorname{mod} 2^k)}} \sum_{\substack{x^{1/3}<p\leqslant x/(qr) \\ p\equiv -r (\operatorname{mod} 2^k)}} 1.$$

Then we apply the prime number (in arithmetic progressions) theorem to write

$$\begin{aligned}
\sum_{\substack{n\leqslant x\\ n\in\mathscr{C}_3(k)}} 1 &\gg \frac{x}{2^k\log x}\sum_{r\leqslant x^{1/6}}\sum_{r<q\leqslant x^{1/3},\, q\equiv r(\mathrm{mod}\, 2^k)}\frac{1}{qr}\\
&\gg \frac{x}{2^{2k}\log x}\sum_{r\leqslant x^{1/6}}\frac{1}{r}\big\{\log\log x^{1/3}-\log\log r+O(1)\big\}\\
&\gg \frac{x}{2^{2k}\log x}\bigg\{(\log\log x^{1/3})(\log\log x^{1/6})-\frac{1}{2}(\log\log x^{1/6})^2+O(\log\log x)\bigg\}\\
&\gg \frac{1}{2^{2k}}\frac{x}{\log x}(\log\log x)^2\\
&\gg \frac{1}{2^{2k}}A_3(x).
\end{aligned}$$

This completes the proof. □

Now we are ready to prove Theorem 6.

For any $\alpha\in(0,1]$, we can choose k and ε such that $(1+\varepsilon)2^{-k}<\alpha\leqslant 2^{-k+1}$. In particular for $\alpha\in(\frac{1}{2},1]$, we can take $k=1$. Thus

$$\begin{aligned}
\sum_{\substack{n\leqslant x,\, n\in\mathscr{A}_3\\ P(w(n))\leqslant\alpha P(n)}} 1 \geqslant \sum_{\substack{n\leqslant x,\, n\in\mathscr{C}_3(k)\\ P(w(n))\leqslant\alpha P(n)}} 1
\geqslant \sum_{\substack{n\leqslant x,\, n\in\mathscr{C}_3(k)\\ P(w(n))\leqslant(1+\varepsilon)2^{-k}P(n)}} 1.
\end{aligned}$$

The asymptotic formula (1.12) follows from (6.4) with $k=1$, (6.5), (6.1) and the trivial inclusion relation $\mathscr{C}_3\subset\mathscr{A}_3$. The inequality (1.13) can be obtained in the same way (i.e. by replacing (6.5) with (6.6)).

7 A heuristic argument on Problem 11

The aim of this section is to present an heuristic proof of

$$\sum_{p\leqslant x,\, P(p+2)=q} 1\to\infty\qquad(x\to\infty), \tag{7.1}$$

where q is a large fixed prime number.

Since $P(n+2)=q$ is equivalent to $n+2=q\ell$ with $P(\ell)\leqslant q$, we have

$$\sum_{n\leqslant x,\, P(n+2)=q} 1=\sum_{\ell\leqslant(x+2)/q,\, P(\ell)\leqslant q} 1=\Psi((x+2)/q,q),$$

where

$$\Psi(x,y) := \sum_{n\leqslant x,\, P(n)\leqslant y} 1.$$

According to [8, Theorem III.5.2], we have

$$\log \Psi((x+2)/q, q) = Z\left\{1 + O\left(\frac{1}{\log q} + \frac{1}{\log\log((x+2)/q)}\right)\right\}$$

uniformly for $x+2 \geqslant q^2 \geqslant 10000$, where

$$Z := \frac{\log((x+2)/q)}{\log q}\log\left(1+\frac{q}{\log((x+2)/q)}\right) + \frac{q}{\log q}\log\left(1+\frac{\log((x+2)/q)}{q}\right).$$

From this we deduce

$$Z \geqslant \frac{q}{2\log q}\log\log x$$

and

$$\Psi((x+2)/q, q) \geqslant (\log x)^{q/(4\log q)} \tag{7.2}$$

for $x \geqslant x_0(q)$, provided the constants q and $x_0(q)$ are suitably large.

According to Cramér's model [9, Section 3.2] it seems reasonable to assume that

$$\sum_{p\leqslant x,\, P(p+2)=q} 1 \approx \sum_{n\leqslant x,\, P(n+2)=q} \frac{1}{\log n} \qquad (x\to\infty). \tag{7.3}$$

Thus (7.2) implies (under the hypothesis (7.3))

$$\sum_{p\leqslant x,\, P(p+2)=q} 1 \gg (\log x)^{q/(4\log q)-1} \to \infty,$$

as $x\to\infty$.

References

[1] R. C. Baker & G. Harman, *Shifted primes without large prime factors*, Acta Arith. **83** (1998), no. 4, 331–361.

[2] Y.-G. Chen & Y. Shi, *Dynamics of the w function and the Green-Tao theorem on arithmetic progressions in the primes*, Proc. Amer. Math. Soc. **136** (2008), 2351–2357.

[3] Y.-G. Chen & Y. Shi, *Distribution of primes and dynamics of the w function*, J. Number Theory **128** (2008), 2085–2090.

[4] W. Goldring, *Dynamics of the w function and primes*, J. Number Theory **119** (2006), 86–98.

[5] B. Green & T. Tao, *The primes contain arbitrarily long arithmetic progressions*, Ann. Math. **167** (2008), 481–547.

[6] C.-H. Jia, *On the inverse problem relative to dynamics of the w function*, Science in China, Ser. A, **52** (2009), no. 5, 849–956.

[7] H. L. Montgomery, *Topics in multiplicative number theory*, Lecture Notes in Math., Vol. 227, Springer, Berlin, 1971.

[8] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, Cambridge Studies in Advanced Mathematics **46**, Cambridge University Press, 1995.

[9] G. Tenenbaum & M. Mendès France, *The prime numbers and their distribution*, Translated from the 1997 French original by Philip G. Spain. Student Mathematical Library, 6. American Mathematical Society, Providence, RI, 2000. xx+115 pp.